AD_____

Award Number: DAMD17-98-1-8168

TITLE: Phage Display Breast Carcinoma cDNA Libraries: Isolation of Clones Which Specifically Bind to Membrane Glycoproteins, Mucins, and Endothelial Cell Surface

PRINCIPAL INVESTIGATOR: Fumiichiro Yamamoto, Ph.D.

CONTRACTING ORGANIZATION: The Burnham Institute
La Jolla, California 92037

REPORT DATE: July 2000

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 2000	Final (1 Jul 98 - 30 Jun 00)

4. TITLE AND SUBTITLE
Phage Display Breast Carcinoma cDNA Libraries: Isolation of Clones Which Specifically Bind to Membrane Glycoproteins, Mucins, and Endothelial Cell Surface

5. FUNDING NUMBERS
DAMD17-98-1-8168

6. AUTHOR(S)
Fumiichiro Yamamoto, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Burnham Institute
La Jolla, California 92037

E-MAIL:
fyamamoto@burnham-inst.org

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 Words)*

We proposed to identify, by using a novel phage display technique, carbohydrate-binding proteins (lectins), which may play an important role in progression and metastasis of breast cancer. Using blood-group H-expressing glycoprotein fraction as bait, we observed enrichment of phage clones expressing sequences from galectin-3, a lectin with an affinity with the blood-group substance. Since we obtained a multiple number of different clones encoding this protein, the binding was selective and real. We then constructed additional libraries including MCF-7 human breast carcinoma and EJG bovine capillary endothelial cell libraries. We performed *in vitro* biopanning experiments using as bait purified fractions of glycoproteins and glycolipids, as well as antibodies against plant lectins, and obtained several dozens of candidate phage clones. However, none of these candidates contained cDNA sequences encoding the same proteins. *In vivo* biopanning was also performed using cultured EJG cells as bait. PCR amplification of the inserts from the phage lysates produced a smear of fragments rather than discrete bands even after the fifth/sixth round of selections. *In situ* biopanning using breast tissue sections as bait was also performed. However, again none of the candidates contained cDNA sequences encoding the same proteins.

14. SUBJECT TERMS
Breast Cancer, Phage display libraries, cDNA cloning, surface binding, carbohydrate-binding proteins (lectins)

15. NUMBER OF PAGES
22

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

___ Where copyrighted material is quoted, permission has been obtained to use such material.

___ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

___ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

N/A In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

X For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

N/A In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

N/A In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

N/A In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Francisco Yamamoto 3/24/00

PI - Signature Date

Table of Contents

Cover.....1

SF 298.....2

Foreword.....3

Table of Contents.....4

Introduction.....5

Body.....6-9

Key Research Accomplishments.....9

Reportable Outcomes.....9

Conclusions.....9

Bibliography.....9

List of Personnel.....9

References.....10-11

Appendices.....12-22

INTRODUCTION

One major function of the carbohydrates in glycoproteins and glycolipids is to present information in the form of three-dimensional structures. The information in oligosaccharide structures is decoded through the binding of carbohydrate recognizing proteins called lectins. Involvement of lectins in diverse biological phenomena, from intracellular routing of glycoproteins to cell-cell adhesion and phagocytosis, has been reported (1). Especially significant is the progress made in elucidating the role of selectins, a subgroup of C-type lectins, in inflammation (2-4). Under physiological flow conditions, leukocytes interact with activated vascular endothelium through leukocyte rolling, followed by firm adhesion, spreading, and extravasation. Selectins, which bind to oligosaccharide structures called sialyl Le^x and sialyl Le^a, were shown to participate in the rolling process. Since both leukocytes and tumor cells, in particular carcinomas and leukemias, express these oligosaccharides on their cell surface, it was postulated that tumor cells might use selectin-carbohydrate interaction during metastasis. Although limited, existing experimental data support this hypothesis. Carbohydrate-protein interactions are not restricted to selectins. Other lectins may be used by tumor cells when interacting with carbohydrates, and these lectins may also play an important role in progression and metastasis. A significant difference was observed in the degree of glycosylation of the mucin, as well as in the N-glycosylation pattern of glycoproteins between normal and cancer breast cells (5-8). Therefore, to explore that possibility breast cancer was chosen as the prime target of our study.

The primary goal of this proposal is to identify lectins that are either produced by breast cancer cells or that are capable of binding to cell surface carbohydrate structures of the breast cancer cells. In order to identify these lectins we chose to employ a powerful technique known as phage display (9-12). Use of the recently introduced T7 phage system allowed us to express proteins up to 1200 amino acids, which might be sufficient to contain the carbohydrate-recognition domain (CRD) of lectins.

BODY

Task 0. *To prove the feasibility of phage display strategy in cloning cDNA encoding proteins with carbohydrate-binding property*

From our experiments with a HeLa cell phage display cDNA library and an isolated blood group H-specific glycoprotein fraction, we obtained results suggesting the possible utility of the phage display strategy in identifying carbohydrate binding proteins *in vitro*. By using this glycoprotein fraction as bait, amplification from the library of multiple phage clones expressing sequences (13-16) from galectin-3 was observed. Galectin-3 is a member of S-type lectins (17-19), and an affinity with the blood-group substance has been reported (19,20). Before expanding our research to breast cancer, we proved that amplification of galectin-3 phage display clones by *in vitro* biopanning was not incidental but rather selective. Southern hybridization, as well as, plaque hybridization was used to prove the selectivity. Results from this portion of work have been published (Yamamoto, M., Kominato, Y., and Yamamoto, F. (1999). Phage display cDNA cloning of protein with carbohydrate affinity. Biochem. Biophys. Res. Commu. 255: 194-199), and the reprint is located in the Appendix.

Task 1. To obtain phage display cDNA clones which specifically bind to isolated glycoproteins and cell surfaces

**Construction of phage display cDNA libraries*

We have constructed a total of 11 phage display cDNA libraries in T7Select 1-1 vector where eight of these libraries contained more than one million independent clones. The size and titers of the amplified libraries were determined and listed in Table 1.

** In vitro biopanning to isolate clones from phage display cDNA libraries*

In addition to the blood group glycoprotein fraction from gastric mucosa, we used commercially available purified glycoprotein and glycoprotein fractions for *in vitro* biopanning experiments. We also used antibodies against plant lectins as bait.

The list of glycolipids, glycoproteins, and antibodies used for *in vitro* biopanning are as follows.

Glycoproteins:
Blood group H-expressing mucin fraction from gastric mucosa, α1-Acid Glycoprotein from human serum, Fibrinogen from human plasma, extracellular matrix glycoprotein fraction of Zonae pellucida from porcine ovary, and sialyl Lewis x-BSA

Glycolipids:
Trihexosylceramide (CTH), Asialoganglioside-GM1, Asialoganglioside-GM2, Monosialoganglioside-GM1, Monosialoganglioside-GM2, and Monosialoganglioside-GM3

Antibodies:

Biotinylated anti-Phaseolus vulgaris agglutinin, biotinylated anti-Concanavalin A (Con A), biotinylated anti-Peanut agglutinin (PNA), and biotinylated anti-wheat germ agglutinin (WGA)

MaxiSorp 96-well plates were coated with glycoproteins and used for binding selections. Some glycolipids were coated onto 96-well PolySorp plates while other charged glycolipids were spotted onto a polyvinylidene difluoride (PVDF) membrane (21). Biopanning of antibodies was performed using streptavidin-coated magnetic beads (Dynal).

* *In vivo* biopanning to isolate clones from phage display *cDNA* libraries

In vivo biopanning experiments were performed using EJG bovine capillary endothelial cells and MCF-7 and BT-20 human mammary carcinoma cells as baits.

* *In situ* biopanning to isolate clones from phage display *cDNA* libraries

We performed *In situ* biopanning using frozen breast tissue sections obtained from the University of Michigan Breast Cancer Tissue Bank.

Task 2. To characterize the selected phage populations as well as isolated individual clones

*Monitoring the enrichment of certain phage clones by PCR using lysates from selected population of phages, determining nucleotide sequences cDNA inserts, and identifying the clones of interest

In vitro biopanning

After the fourth or fifth round of *in vitro* biopanning, PCR amplification of cDNA inserts was performed and the PCR products were analyzed by agarose gel electrophoresis. Several discrete bands were observed from each selection, and the selections were repeated using this variety of glycoproteins, glycolipids, and antibodies as bait. These bands were excised from the gel, DNA was gel-purified, and the nucleotide sequences of the cDNA inserts were determined directly or after cloning into a plasmid vector. The nucleotide sequences of 153 bands were determined (Table 2). Homology search was performed using the BLAST program. Unexpectedly, however, none of the phages contained the cDNA sequences encoding the same proteins except that two identical clones encoding benzodiazapine receptor were obtained using two different antibodies against plant lectins as bait.

In vivo biopanning

Different from the above-mentioned *in vitro* biopanning experiments, no discrete bands were observed after the sixth round of selection. Apparently many phage clones bound to the cell surface (and/or cell culture plate substratum). We next examined whether phages could be further screened based on stimulatory/inhibitory activity on cell growth. For this purpose, 6x48 clones (may not be completely independent) were randomly selected from the fifth/sixth rounds of EJG cell selections and infected to the host bacteria. Phage lysates were then added onto the EJG cell culture in 24-well dishes in duplicate. After 72 hours of incubation, cell growth was monitored by MTT assay (22).

Three clones each showing the highest or the lowest values in duplicate MTT experiments were selected out of the respective set of 48 clones. Therefore, a total of 36 clones were analyzed in the second round of MTT assays. We anticipated that those phage clones would exhibit the same tendency as observed in the first round forming two separate populations (one with stimulatory activity and the other with inhibitory activity on EJG cell growth). Disappointing results were obtained. Some of the clones previously identified to enhance the cell growth in the first round of selection turned out to repress the growth in the second round, and vice versa. We went on the next screening hoping to identify several true candidates out of incidentally selected clones, but the third selection was not successful either. Again, no separation between growth stimulatory clones and growth inhibitory clones were observed. Results were random and not reproducible. Since the phage titers were approximately $5x10^7$ pfu/μl and 500 μl of lysate was added to a 2ml of cell culture, the concentrations of the fusion proteins ($5x10^7$ x 500 molecules/2.5ml = 10^{13} molecules/l = $10^{13}/6.02x10^{23}$ mol/l = 17 $x10^{-12}$ mol/l = 17pM) may have not been sufficient to induce growth changes. In order to achieve higher concentration of fusion proteins, transfer of the cDNA inserts from phage display vector to a eukaryotic expression vector may be necessary. We sequenced 72 phage clones from the population selected with EJG cells as bait, and they are listed in Table 2.

In situ biopanning
After the fifth round of biopanning, several discrete bands were observed. We plated phages at low density and PCR-amplify cDNA inserts from individual plaques. Nucleotide sequences of 20 inserts were determined. Except that 2 sequences were from the same clone, no multiple clones encoding the same protein were observed.

Task 3. To characterize the identified genes and their gene expression
*Performing Northern hybridization and/or RT-PCR to determine the mRNA level using RNA from normal and tumor breast cell lines, and performing Southern hybridization using DNA from normal and tumor breast cell lines

We have not performed these experiments since no true candidates were obtained from Task 2. We obtained many interesting candidate clones, however, we were unable to obtain a multiple number of different clones encoding the same protein. In addition, many clones encoded relatively abundant proteins. These included glutaraldehyde 3 phosphate dehydrogenase and ribosomal proteins. Further, the directions and coding frames of the cDNA sequences varied. These results suggested that these candidates might have been technical artifacts rather than true candidates. Therefore, we tested that possibility. We used purified glycoproteins as baits and the previously selected secondary or tertiary round of phage lysates instead of cDNA libraries. Complementary DNA libraries contained millions of different clones, whereas those lysates contained hundreds (or thousands) of clones and some clones were already enriched. Therefore, the same clones that were enriched after the fourth or fifth round of biopanning starting from library were expected to enrich at the earlier round. The results showed the absence of selectivity, suggesting opportunistic enrichment. Then we have changed the

binding and washing conditions as well as the method of refreshing phage libaries (23) without success.

KEY RESEARCH ACCOMPLISHMENTS

- Demonstration of the phage display technique to clone cDNA of carbohydrate-binding protein

REPORTABLE OUTCOMES

- A paper reporting the feasibility of phage display approach to clone cDNA encoding carbohydrate-binding protein was published.

CONCLUSIONS

Although we could prove the feasibility of the phage display technique to clone carbohydrate-binding proteins, we have not succeeded in identifying previously unrecognized lectin molecules. *In vitro* and *in situ* biopannings produced many candidates, however, none came from different cDNA clones encoding the same proteins except for galectin-3. Results from *in vivo* biopanning suggested the necessity of functional analysis.

BIBLIOGRAPHY

- **Publications**

M. Yamamoto, Y. Kominato, and F. Yamamoto (1999). Phage display cDNA cloning of protein with carbohydrate affinity. Biochem. Biophys. Res. Commu. 255: 194-199

LIST OF PERSONNEL

- Fumiichiro Yamamoto, Ph.D., PI
- Miyako Yamamoto, B.S., Research Technician

REFERENCES

1. Drickamer, K. Molecular structure of animal lectins. in *Molecular Glycobiology* (eds. Fukuda, M. & Hindsgual, O.) 53-87 (IRL Press, Oxford, UK, 1994).

2. Phillips, M.L. *et al.* ELAM-1 mediates cell adhesion by recognition of a carbohydrate ligand, sialyl-Lex. *Science* **250**, 1130-1132 (1990).

3. Lowe, J.B. *et al.* ELAM-1--dependent cell adhesion to vascular endothelium determined by a transfected human fucosyltransferase cDNA. *Cell* **63**, 475-484 (1990).

4. Walz, G., Aruffo, A., Kolanus, W., Bevilacqua, M. & Seed, B. Recognition by ELAM-1 of the sialyl-Lex determinant on myeloid and tumor cells. *Science* **250**, 1132-1135 (1990).

5. Ceriani, R.L., Peterson, J.A., Blank, E.W. & Lamport, D.T. Epitope expression on the breast epithelial mucin. *Breast Cancer Res Treat* **24**, 103-113 (1993).

6. Hanisch, F.G., Stadie, T.R., Deutzmann, F. & Peter-Katalinic, J. MUC1 glycoforms in breast cancer--cell line T47D as a model for carcinoma-associated alterations of 0-glycosylation. *Eur J Biochem* **236**, 318-327 (1996).

7. Sikut, R., Zhang, K., Baeckstrom, D. & Hansson, G.C. Distinct sub-populations of carcinoma-associated MUC1 mucins as detected by the monoclonal antibody 9H8 and antibodies against the sialyl-Lewis a and sialyl-Lewis x epitopes in the circulation of breast-cancer patients. *Int J Cancer* **66**, 617-623 (1996).

8. Lloyd, K.O., Burchell, J., Kudryashov, V., Yin, B.W.T. & Taylor-Papadimitriou, J. Comparison of O-linked carbohydrate chains in MUC-1 mucin from normal breast epithelial cell lines and breast carcinoma cell lines. Demonstration of simpler and fewer glycan chains in tumor cells. *J Biol Chem* **271**, 33325-33334 (1996).

9. *Phage Display of Peptides and Proteins: A Laboratory Manual*, (Academic Press, San Diego, 1996).

10. Smith, G.P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315-1317 (1985).

11. Winter, G. & Milstein, C. Man-made antibodies. *Nature* **349**, 293-299 (1991).

12. Clackson, T., Hoogenboom, H.R., Griffiths, A.D. & Winter, G. Making antibody fragments using phage display libraries. *Nature* **352**, 624-628 (1991).

13. Cherayil, B.J., Chaitovitz, S., Wong, C. & Pillai, S. Molecular cloning of a human macrophage lectin specific for galactose. *Proc Natl Acad Sci U S A* **87**, 7324-7328 (1990).

14. Robertson, M.W., Albrandt, K., Keller, D. & Liu, F.T. Human IgE-binding protein: a soluble lectin exhibiting a highly conserved interspecies sequence and differential recognition of IgE glycoforms. *Biochemistry* **29**, 8093-8100 (1990).

15. Oda, Y., Leffler, H., Sakakura, Y., Kasai, K. & Barondes, S.H. Human breast carcinoma cDNA encoding a galactoside-binding lectin homologous to mouse Mac-2 antigen. *Gene* **99**, 279-283 (1991).

16. Raz, A. *et al.* Molecular cloning and chromosomal mapping of a human galactoside-binding protein. *Cancer Res* **51**, 2173-2178 (1991).

17. Albrandt, K., Orida, N.K. & Liu, F.T. An IgE-binding protein with a distinctive repetitive sequence and homology with an IgG receptor. *Proc Natl Acad Sci U S A* **84**, 6859-6863 (1987).

18. Abbott, W.M., Hounsell, E.F. & Feizi, T. Further studies of oligosaccharide recognition by the soluble 13 kDa lectin of bovine heart muscle. Ability to accommodate the blood-group-H and -B-related sequences. *Biochem J* **252**, 283-287 (1988).

19. Leffler, H. & Barondes, S.H. Specificity of binding of three soluble rat lung lectins to substituted and unsubstituted mammalian beta-galactosides. *J Biol Chem* **261**, 10119-10126 (1986).

20. Sato, S. & Hughes, R.C. Binding specificity of a baby hamster kidney lectin for H type I and II chains, polylactosamine glycans, and appropriately glycosylated forms of laminin and fibronectin. *J Biol Chem* **267**, 6983-6990 (1992).

21. Taki, T. & Ishikawa, D. TLC blotting: application to microscale analysis of lipids and as a new approach to lipid-protein interaction. *Anal Biochem* **251**, 135-143 (1997).

22. Carmichael, J., DeGraff, W.G., Gazdar, A.F., Minna, J.D. & Mitchell, J.B. Evaluation of a tetrazolium-based semiautomated colorimetric assay: assessment of radiosensitivity. *Cancer Res* **47**, 943-946 (1987).

23. Zozulya, S., Lioubin, M., Hill, R.J., Abram, C. & Gishizky, M.L. Mapping signal transduction pathways by phage display. *Nat Biotechnol* **17**, 1193-1198 (1999).

APPENDIX

Table 1. List of phage display cDNA libraries

MCF-7 Breast adenocarcinoma	1.06×10^5 clones	4×10^7 pfu/μl
MCF-7 Breast adenocarcinoma	1.26×10^7 clones	7.23×10^7 pfu/μl
PC-3 prostate adenocarcinoma	1.94×10^5 clones	4.4×10^7 pfu/μl
PC-3 prostate adenocarcinoma	1.63×10^7 clones	2.80×10^7 pfu/μl
HeLa cell transfectant	1.77×10^6 clones	3.9×10^7 pfu/μl
HeLa cell transfectant	3.05×10^6 clones	3.25×10^7 pfu/μl
HeLa cell transfectant	5.29×10^6 clones	3.44×10^7 pfu/μl
EJG bovine capillary endothelial cell	1.54×10^7 clones	2.80×10^7 pfu/μl
Human liver	4.68×10^5 clones	5×10^7 pfu/μl
Human liver	2.55×10^7 clones	3.48×10^7 pfu/μl
Human lymph node	4.63×10^6 clones	5.27×10^7 pfu/μl

The number of the independent clones and the titers of the amplified libraries are shown.

Table 2. Sequence homology of the phage clones obtained by selections

(I). In vitro selection

A. Glycoproteins

1. Blood group H-expressing mucin fraction

M	X89718	1.9e-16	Putative 26S protease subunit mRNA
M	AA134606	1.1e-88	EST
M	X98614	2e-92	cytokeratin 8 mRNA
M	AC002558	e-114	chr. 17 identical
M	AA627181	e-108	EST
M	X06617	e-128	ribosomal protein S11 mRNA
M	AI394640	e-153	EST
M	J05176	e-113	alpha 1-antichymotrypsin mRNA
M	X90826	e-170	USF2 mRNA
M	AA040063	e-162	EST
M	D38112	6e-32	mitochondrial DNA
MPELiLy	Y13710	0.0	alternative activated macrophage-specific CC cytokine mRNA
MPELiLy	T89015	2e-19	KIAA 0220 mRNA
MPELiLy	AA225548	3e-30	EST/Alu
MPELiLy	S416941	6e-4	PTMAP2 mRNA
MPELiLy	L41066	7e-53	NF-AT3 mRNA
MPELiLy	U90426	0.0	nuclear RNA helicase mRNA
MPELiLy	AA83124	3e-15	Alu
MPELiLy	AI315224	2.2	
MPELiLy	AA832124	3e-15	Alu
MPELiLy	Z54072	e-17	Alu
MPELiLy	AC004383	0.24	
MPELiLy	M36682	e-36	laminin-binding protein (galectin-3) mRNA
MPELiLy	AI253442	6e-50	cytochrome C mRNA
MPELiLy	AC002487	4.9	
MPELiLy	AC004896	3e-12	Alu
MPELiLy	H75348	2e-25	
MPELiLy	AA71349	4e-119	EST
MPELiLy	AC004911	2.1	
MPELiLy	U63090	3e-67	alpha 2,3 sialyltransferase mRNA
MPELiLy	W33034	2e-99	EST
MPELiLy	R76396	4e-23	Alu
MPELiLy	AA804821	e-60	ATP synthetase mRNA
MPELiLy	AC005224	3e-46	Alu
MPELiLy	AA384311	e-13	EST
MPELiLy	AF042506	4e-91	cytochrome b mRNA
MPELiLy	AA327241	0.008	

MPELiLy	Y00097	3e-36	p68 protein mRNA
MPELiLy	AC00310	6e-38	Alu
MPELiLy	X75362	0.014	
MPELiLy	AA687598	6e-54	EST
MPELiLy	S80305	7e-29	beta2 glycoprotein I mRNA
MPELiLy	U24576	7e-84	breast tumor autoantigen mRNA
MPELiLy	X02308	e-136	thymidylate synthetase mRNA
MPELiLy	U35451	0.0	heterochromatin protein p25 mRNA
MPELiLy	X93334	0.0	mitochondrial DNA
MPELiLy	AC003034	e-120	BAC clone
MPELiLy	Z84480	0.002	
2. Sialyl Le^x-BSA			
MPELiLy	AF069378.1	3e-82	Insulin-like growth factor receptor II
MPELiLy	Z97353.3	3e-55	ribosomal protein L10 mRNA
MPELiLy	L06237	5e-9	microtubule-associated protein 1B mRNA
MPELiLy	NM 001306.1	4e-60	Claudin 3 mRNA
3. Zonae pellucida glycoprotein fraction			
MPELiLy	U93305	6e-10	A4 differentiation dependent triple LIM domain protein mRNA
MPELiLy	NM 000970.1	e-112	ribosomal protein L27a mRNA
4. Fibrinogen			
MPELiLy	U63290	0.54	
5. alpha 1 Acid glycoprotein fraction			
MPELiLy	AA996393	5e-7	EST

B. Glycolipids

1. CTH			
MPELiLy	AJ000039	3e-93	G3PDH
MPELiLy	AF004338	e-147	16S rRNA
MPELiLy	Y11395	0.0	p40 protein mRNA
MPELiLy	D38524	e-168	5' nucleotidase mRNA
MPELiLy	AJ130972	0.0	U2 sn RNP-specific A protein mRNA
2. Asialo GM2			
MPELiLy	AJ130972	e-96	U2 sn RNP mRNA
MPELiLy	AB012622	3e-97	Tapasin mRNA
MPELiLy	AF055004	0.0	KIAA0763 gene mRNA
MPELiLy	X04225	e-112	HCC alpha 1 microglobulin mRNA
MPELiLy	Z84480	0.002	
3. Asialo GM1			
MPELiLy	X86343	e-157	complement C7 mRNA
MPELiLy	AB012622	3e-97	Tapasim mRNA
MPELiLy	D88984	e-132	Ampd 3 gene
MPELiLy	S74678	e-123	heterogeneous nuclear RNP complex K mRNA
MPELiLy	M29064	5e-06	hnRNP B1 protein mRNA
MPELiLy	Z75894	0.14	
4. GM3			
MPELiLy	NM 003191.1	4e-59	thr-tRNA synthetase mRNA
MPELiLy	R78124	1,7	
MPELiLy	X85133	3e-49	p53 associated RBQ-1 mRNA
MPELiLy	AJ000039	7e-40	G3PDH mRNA
5. GM2			
MPELiLy	U96922	3e-93	inositol polyphosphate 4-phosphatase type II mRNA
MPELiLy	U68758	0.0	pyroline-5-carboxylate synthetase
MPELiLy	AF035429	2e-97	cytochrome oxidase subunit I mRNA
MPELiLy	AI093867	e-144	Hox A10 mRNA
MPELiLy	Y11395	0.0	p40 protein mRNA
MPELiLy	U011184	e-112	flightless gene product mRNA
MPELiLy	X56932	2e-66	23kD highly basic protein mRNA
MPELiLy	AC005303	0.11	
MPELiLy	M11233	e-112	cathepsin O mRNA
6. GM1			
MPELiLy	X85786	0.0	DNA binding regulatory factor mRNA
MPELiLy	L05095	0.0	ribosomal protein L30 mRNA
MPELiLy	AA977240	e-176	EST

MPELiLy	D42039	0.0	KIAA 0081 gene mRNA

C. Antibodies

1. anti-Phaseolus vulgaris agglutinin

H	NM00714	e-139	benzodiazapine receptor
H	NM001457	e-148	anti-binding protein 278
H	AL022068.2	0.015	
H	AL006943.25	0.072	
H	NM001527	4e-26	histone deacetylase 2
H	AW058612.1	2e-60	EST
H	NC001807.2	e-142	mitochondria

2. anti-Concanavalin A (Con A)

H	NM001010.1	e-117	ribosomal protein S6
H	NM002687.1	7e-80	desmosome associated pirin
H	NM002032.1	6e-98	ferritin, heavy polypeptide 1
H	NM006371.1	e-132	cartilage associated protein
H	M11167.1	e-106	28S rRNA
H	NM001067.1	e-156	topoisomerase II alpha
H	AC004829.2	4.7	
H	NM002392.1	0.0	double minute 2 p53 binding protein
H	G25247	8e-57	EST
H	U89311.1	e-38	nucleophosmin phosphoprotein

3. anti-Peanut agglutinin (PNA)

H	X92545	2e-6	vigilin
H	AF078845.1	0.0	16.7kDa protein
H	AL136369	0.001	
H	G25835	3e-7	STS
H	AF151080	0.0	HSPC 246 mRNA
H	NM000985.1	e-141	ribosomal protein L17
H	AA480940.1	6e-88	EST
H	NM005566.1	0.0	lactose dehydrogenase A
H	NM005688	e-155	ATP binding protein
MPELiLy	NM003380	0.0	vimentin
MPELiLy	AP000694	5e-92	
MPELiLy	NM000967	0.0	ribosomal protein L3
MPELiLy	NM000062.1	0.0	complement component 1 (C1) inhibitor
MPELiLy	NM000090.1	7e-48	typeIII collagen
MPELiLy	NM002948.1	e-127	ribosomal protein L15
MPELiLy	AW665473.1	1.1	
MPELiLy	NM007315.1	e-113	signal transducer and activator 1 (STAT1)
B	AJ242682	3e-66	ETAA16 protein
B	AL008639.15	0.35	
B	AC009498.3	9e-82	
B	NM001011	5e-77	ribosomal protein S7
B	NM002100	9e-36	hexabrachion (tenascin C)
B	AW572242.1	8e-54	

4. anti-wheat germ agglutinin (WGA)

H	ACC005924.2	3e-84	PAC
H	AC005606	2e-52	genomic
H	NM000714.2	e-148	benzodiazapine receptor
H	AP001040	e-131	genomic
H	AB019438	e-101	Ig heavy chain variable region
H	NM007104	e-114	ribosomal protein L109
H	NM002046	e-150	glutaraldehyde 3 phosphate dehydrogenaseH

(II). In vivo biopanning

A. EJG cells

MPH	U14970	3e-113	ribosomal protein S5 mRNA
MPH	X73974	4e-140	ribosomal protein L4 mRNA
MPH	X01630	3e-171	argininosuccinate synthetase mRNA
MPH	AA722510	5e-133	Kappa casein precursor mRNA
MPH	S35960	4e-155	laminin receptor mRNA
MPH	AA427423	7.1e-65	ribosomal protein S16 mRNA

MPH	R51271	0.18	
MPH	U57344	2e-148	homeobox protein Meis3 mRNA
MPELiLy	Z28663	6e-30	EST
MPELiLy	N58329	7e-40	EST/Alu
MPELiLy	AA523029	4e-72	ribosomal protein L9 mRNA
MPELiLy	H56276	3e-51	carbamoyl-phosphate synthetase mRNA
MPELiLy	W28842	9e-89	EST
MPELiLy	AC005368	e-110	chr.5
MPELiLy	M81310	0.43	
MPELiLy	X56932	0.0	23KD highly basic protein mRNA
MPELiLy	AJ002030	3e-54	putative progesterone binding protein mRNA
MPELiLy	AA418401	e-151	EST
MPELiLy	AC005376	1.7	
MPELiLy	AA720919	3e-14	EST
MPELiLy	L29394	e-164	haptoglobin alpha 2FS beta mRNA
MPELiLy	AC002350	3e-45	Alu
MPELiLy	AC003100	2e-15	
MPELiLy	AA134339	0.08	
MPELiLy	AA195932	0.0	EST
MPELiLy	AA411274	2e-97	
MPELiLy	N58329	9e-42	EST/Alu
MPELiLy	AC002558	0.0	EST
MPELiLy	AC002300	0.088	
MPELiLy	U42386	e-175	fibroblast growth factor mRNA
MPELiLy	AL031228	5e-68	complement gene
MPELiLy	U00947	2e-52	(CAC)n/(GTG)n containing mRNA
MPELiLy	D26488	4e-17	KIAA0007 gene
MPELiLy	M10098	4e-73	18S rRNA
MPELiLy	H41283	6e-07	EST
MPELiLy	AB006198	0.0	SART-1 mRNA
MPELiLy	AA459599	e-173	EST
MPELiLy	U92980	e-113	DT1P1A10 mRNA
MPELiLy	X72889	e-126	hbrm mRNA
MPELiLy	C04745	e-147	EST
MPELiLy	M55670	e-177	protein Z mRNA
MPELiLy	X62996	0.0	mitochondrial DNA
MPELiLy	AF001892	5e-78	MEN1 region epsilon/beta mRNA
MPELiLy	AD000813	3e-6	chr.19
MPELiLy	U66306	e-121	retinoid X receptor alpha mRNA
MPELiLy	H87626	8e-27	EST
MPELiLy	AL009181	0.002	PAC
MPELiLy	J02625	e-165	cytochrome P-450j
MPELiLy	X62153	e-116	P1 protein mRNA
MPELiLy	Z33502	0.007	
MPELiLy	AF038186	8e-95	clone 23914 mRNA
MPELiLy	AB000517	e-141	CDP-diacylglycerol synthetase mRNA
MPELiLy	V00654	2e-47	mitochondrial DNA
MPELiLy	AA381346	3e-5	EST
MPELiLy	U65590	6e-44	IL-1 receptor antagonist IL-1Ra
MPELiLy	U48705	0.009	
MPELiLy	M18981	e-100	prolactin receptor associated protein mRNA
MPELiLy	D55674	e-105	heterogeneous nuclear ribonucleoprotein D mRNA
MPELiLy	L33264	8e-35	CDC2-related protein kinase
MPELiLy	AA405094	e-123	EST
MPELiLy	J01415	0.0	mitochondrial DNA
MPELiLy	X03559	0.0	F1-ATPase beta subunit mRNA
MPELiLy	L06328	e-166	voltage-dependent anion channel protein mRNA
MPELiLy	AB010066	e-34	RBP56/hTAFII68 mRNA
MPELiLy	AL022098	e-100	chr.6
MPELiLy	Z69723	0.16	
MPELiLy	AF001542	2e-53	EST
MPELiLy	AF03694	0.45	
MPELiLy	AA964847	1.5	
MPELiLy	AC000114	8e-56	chr. Xq23

MPELiLy	U85088	e-8	coagulation factor XI gene
MPELiLy	AC005952	5e-8	chr. 19/Alu

(III). In situ biopanning

Breast cancer tissue section

MPELiLy	X75349	e-108	clotting factor IX
MPELiLy	U62531	0.0	AE2 anion exchanger
MPELiLy	NM000994	0.0	ribosomal protein L32
MPELiLy	AC005998	e-142	L1 repeat
MPELiLy	NM000969	6e-63	ribosomal protein L5
MPELiLy	AI594532.1	0.03	
MPELiLy	NM000064	0.0	complement component 3 (C3)
MPELiLy	AW014647	e-94	
MPELiLy	NM000040	6e-99	apolipoprotein III
MPELiLy	R72920	8e-59	EST
MPELiLy	NM005032	e-128	plastin 3
MPELiLy	AL034418	0.0	glucosamine-6-sulfatase
MPELiLy	AF159295	e-106	18S rRNA
MPELiLy	NM000224	0.0	cytokeratin 18
MPELiLy	NM002215	0.0	inter alpha globulin inhibitor
MPELiLy	AW105663	7e-47	EST
MPELiLy	U712131	e-133	glutathione S-transferase

Abbreviation for the phage display cDNA libraries are as follows. M: MCF-7, P: PC-3, H: HeLa transfectant, E: EJG, Li: Liver, Ly: Lymph node, and B: breast tumor. Except the breast tissue (B) cDNA library which was constructed in T7Select10-3 vector, all the other libraries were constructed in T7Select 1-1b vector. Frame of the codons and orientation of the cDNA inserts varied. Therefore, it should be noted that homology in the nucleotide sequence does not necessarily implicate homology in the amino acid sequence.

transferase transfers GalNAc (*N*-acetyl-D-galactosamine) residue to the substrate H by alpha 1→3 glycosidic linkage, whereas B transferase transfers D-galactose. We used the A transferase as a model enzyme, and examined the feasibility of cDNA cloning of glycosyltransferases by a novel technique employing phage display.

The phage display technology is based on the surface expression of the peptide sequences fused with phage capsid protein. Until recently, most applications were restricted to clone phage particles expressing variable domains of antibodies specific to certain antigens (23–25). Phage display *peptide* libraries made with synthetic oligonucleotides have recently been utilized to identify peptide sequences that interact with a variety of bait ligands, such as proteins, peptides, DNAs, RNAs, and oligonucleotides (26, 27). In addition to the *in vitro* selections, their applications have been extended to *in vivo* selections (28, 29). Surface display was also used for epitope mapping (30). Coat protein gene III of filamentous phage has been commonly used for phage display, but other fusion vectors such as flagellin gene (fliC) of *E. coli* and capsid protein gene 10 of T7 phage have become available. While the C-terminus of the peptide sequence is fused with gene III capsid protein in the filamentous phage system (N-terminus of the peptide is linked to pelB leader sequence), the N-terminus of peptide sequence fuses with gene 10 capsid protein in the T7 system. With this T7 system, it is now possible to produce the phages that display protein sequences up to 1200 amino acid residues long.

We constructed a phage display *cDNA* library using RNA from cells that stably expressed A transferase, and performed biopanning experiments using, as a bait ligand, crude mucin fraction containing blood group H-specific glycoproteins. So far no enrichment of the phages that expressed A transferase fusion protein has been obtained. Unexpectedly, however, selective augmentation was observed of the phages that expressed the fusion proteins with galectin-3, a soluble β-galactoside-binding (S-type) lectin. Because of this lectin's known affinity with the blood group-specific oligosaccharides (13, 14), these results demonstrated, for the first time, that the phage display was successful in cloning cDNAs encoding a protein with binding capacity to carbohydrates.

EXPERIMENTAL PROCEDURES

Construction of a phage display cDNA library. We constructed a phage display *cDNA* library in bacteriophage T7Select System. HeLa cell transformant that stably expressed large amounts of A antigens was used for an RNA source. Total RNA was prepared with RNA STAT-60 reagents (TEL-TEST, Friendswood, TX), and poly A+ RNA was selected by oligo-dT affinity chromatography. Directional cDNA library was constructed by using Directional RH Random Primer cDNA Synthesis Kit and T7Select 1-1b Cloning Kit (Novagen, Madison WI). After *in vitro* packaging, a small aliquot was used to titrate the library. The remaining library was amplified prior to biopanning. Amplified library was also titrated.

Crude extraction of mucins. The human stomach mucosa of blood group O secretor individuals were used to prepare crude mucin fraction that contained blood group H-specific glycoproteins (31). Mucosa was homogenized, digested with pepsin, and ethanol precipitated. After phenol extraction, soluble fraction was precipitated with 10% ethanol.

Biopanning protocol. The MaxiSorp 96 well plate (Nunc, Rochester, NY) was coated for two hours at room temperature with 50 μl of PBS solution containing 5 ng, 50 ng, 500 ng, 5 μg, or 50 μg of crude mucin. After the blocking with PBS containing 3% BSA overnight at room temperature, the wells were washed with PBS. Next, the glycoprotein-coated wells were incubated for 2 hours with 5.3×10^8 phages in Tris-buffered saline (TBS) supplemented with 20 mM $MnCl_2$ and 0.05% Tween 20, followed by washing with TBS. Then the overnight culture of BLT 5403 strain of *E. coli* was added, the wells were incubated for an hour, the contents were transferred into L-Broth supplemented with 50 μg/ml carbenicillin and additional overnight culture of BLT 5403, and the phages were grown until lysis occurred. Phage lysate was then diluted. The entire procedures were repeated two more times.

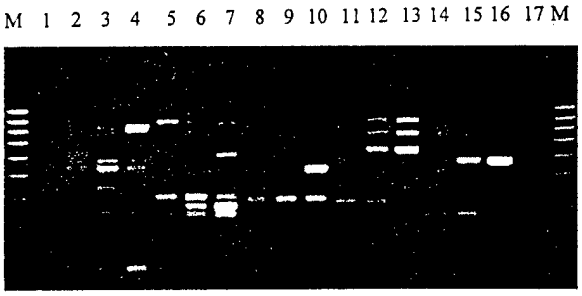
PCR amplification and DNA sequencing. To determine the enrichment effect of biopanning, phage lysates containing selected populations of phages from each round of biopanning were used as templates for PCR amplification. T7 UP and DOWN primers (Novagen) corresponding to the sequences in the phage vector DNA that neighbor the inserts were used to amplify cDNA sequences. PCR was performed in 20 μl of a reaction mixture containing 1× PCR buffer (10 mM Tris–HCl (pH 8.3), 50 mM KCl, 2.5 mM $MgCl_2$, 0.001% Gelatin), 10% DMSO, 250 μM dNTP mix, 40 pmol each of primers, and 1 units of Taq DNA polymerase (BRL-Life Technologies, Gaithersburg, MD). Amplified fragments were analyzed through 3% agarose gel electrophoresis, gel purified with GeneClean II Kit (BIO 101, Vista, CA), and subjected to DNA sequencing using PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (PE-ABI, Foster City, CA) directly and after T-A cloning into the plasmid vector, pT-Adv (Clontech, Palo Alto, CA).

Probe preparation and hybridization. Plasmid DNA containing the galectin-3 cDNA sequence was digested with *Eco*RI. The insert DNA was gel-purified after 3% agarose gel electrophoresis. Random hexamer labeling method was employed to radiolabel the fragment using ^{32}P-dCTP (NEN-Dupont, Boston, MA) and Random Primed DNA Labeling Kit (Boehringer-Mannheim, Indianapolis, IN). PCR-amplified DNA from the selected populations of phages analyzed through agarose gel electrophoresis as above was denatured, neutralized, Southern-transferred onto a Nylon membrane, and UV cross-linked. Approximately 500 phages were plated from each selection, plaque-transferred onto Nylon membranes and processed. These membrane filters were hybridized overnight with the galectin-3 cDNA probe in Pipes buffer (40 mM Pipes (pH 6.4), 1 mM EDTA (pH 8.0). 0.4 M NaCl) with 80% formamide and 1% SDS at 42°C. The filters were then washed and exposed to the X-ray film.

RESULTS

We constructed a HeLa cell transformant phage display *cDNA* library in T7Select 1-1b vector. Most, if not all, of the glycosyltransferases, including A transferase, possess their catalytic domains at the C-terminal side (300–550 amino acid residues depending on the enzymes) of the proteins. Although the transmembrane domain is important for the Golgi localization, several experimental data suggested that

(A)



(B)

FIG. 1. PCR amplification of cDNA inserts from phage populations selected by biopanning. (A). Phage lysates were heat-denatured before being used as templates for PCR. Lysates used are from; original phage display HeLa cell transformant *cDNA* library (lane 1), phage populations obtained after one (lane 2), two (lane 3), and three (lane 4) rounds of selections with 5 ng, or after one (lane 5), two (lane 6), and three (lane 7) rounds of selections with 50 ng, or after one (lane 8), two (lane 9), and three (lane 10) rounds of selections with 500 ng, or after one (lane 11), two (lane 12), and three (lane 13) rounds of selections with 5 μg, or after one (lane 14), two (lane 15), and three (lane 16) rounds of selection s with 50 μg of blood group-specific glycoproteins. Lane 17 shows the result of PCR without lysate as a negative control. M denotes 100-bp DNA ladder marker. (B). DNA was transferred from the above gel onto a Nylon membrane filter, and the filter was hybridized with a ^{32}P-radiolabeled galectin-3 specific probe. A total of five bands were hybridized with the probe.

this signal/anchor domain and the N-terminal cytoplasmic domain were not required for the enzymatic activity (32). T7Select 1-1b display vector can display on the surface of T7 capsid proteins up to 1,200 amino acid residues in size. Target sequences are fused to the C- terminus of the 10B capsid protein. Therefore, this vector was chosen to assess the feasibility of phage display approach to clone glycosyltransferase cDNAs.

For an RNA source, we used HeLa cell transformant that expressed A transferase. The library contained 5.3×10^6 independent phage clones, and the average insert size was 760 bp. Because only 0.1–1 copies of fusion proteins are displayed on T7Select 1-1b phages, one hundred times the size of library (5.3×10^8 phages) was used for the initial round of biopanning selections with blood group glycoproteins, the substrate for A transferase enzymatic reaction, as a target.

We anticipated the enrichment of phages that displayed the A transferase structure through the rounds of selections. Results from the plaque hybridization experiments of the selected populations of phages using a radiolabeled A transferase cDNA probe, however, showed no signs of enrichment (data not shown). Then we examined whether certain populations of phages might be enriched during these biopanning procedures. Using the heat-denatured phage lysate as a template for PCR, we amplified cDNA inserts from the selected populations of phages, and observed enrichment of certain phages as judged by the increased band intensity through the rounds of selections.

Enrichment was observed irrespective of the target concentrations (Fig. 1A). A single major band was amplified from the phage population obtained after three rounds of selections with 50 μg of the target ligand. When 5 μg and 500 ng of targets were used, three and two major bands were amplified, respectively. Enrichment was also seen with the phage populations selected with lesser concentrations of the targets. Nucleotide sequences of these PCR-amplified bands were determined by direct sequencing as well as after cloning into a plasmid vector. Surprisingly, several of these sequences were derived from the cDNA sequence encoding galectin-3, a member of carbohydrate-binding proteins known as lectin. Nucleotide and the deduced amino acid sequences of those bands from galectin-3 cDNA are schematically shown around the fusion junction in Fig. 2. Although the locations and sizes of the inserts varied, their reading frames were identical to that of the fused partner T7 capsid protein. Their C-terminal ends were conserved.

Galectin-3, also known as IgE-binding protein (epsilon-BP), carbohydrate-binding protein 35 (CBP 35), or as macrophage cell-surface protein Mac-2, was originally identified in rat basophilic leukemia cells by virtue of its

Phages

```
                          298
clone 1.  G AAT TCA AGC GCC TAC CCT GCC ACT GGC CCC TAT ...
                          100
                          Ala Tyr Pro Ala Thr Gly Pro Tyr ...

clone 2.  G AAT TCA AGC GCC AGC AGC CGT CCG GAG CCA GCC AAC
                          1
            GAG CGG AAA ATG GCA GAC AAT TTT TCG CTC CAT ...
                          1
                          Met Ala Asp Asn Phe Ser Leu His ...

                          232
clone 3.  G AAT TCA AGC GTC TAC CCA GGG CCA CCC AGC GGC ...
                          78
                          Val Tyr Pro Gly Pro Pro Ser Gly ...

                          346
clone 4.  G AAT TCA AGC GTG CCT TAT AAC CTG CCT TTG CCT ...
                          116
                          Val Pro Tyr Asn Leu Pro Leu Pro ...
```

FIG. 2. Nucleotide and the deduced amino acid sequences of phage clones containing galectin-3 cDNA sequence. Nucleotide and the deduced amino acid sequences from the four phage clones that expressed galectin-3 fusion proteins are schematically shown around the Eco RI fusion junctions with gene 10 B in the T7Select 1-1b Vector. The A residue of initiation codon and the initiation codon are numbered 1 in the nucleotide and the deduced amino acid sequences, respectively. The sequences derived from the Eco RI linker are underlined.

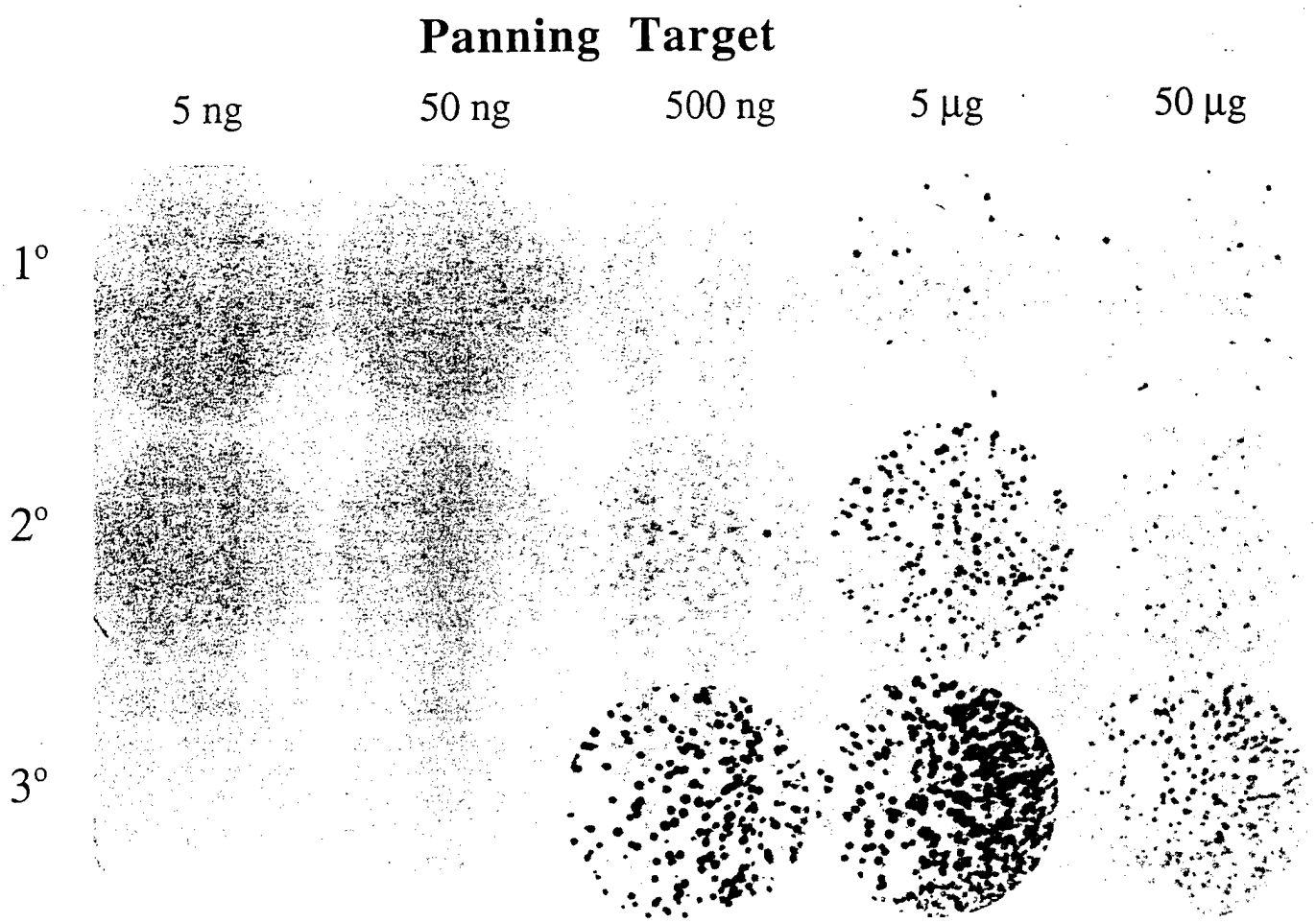


FIG. 3. Hybridization of phage plaques with a galectin-3 probe. Phages from selections were plated at approximately 500 plaques per plate. After plaque-transfer onto Nylon membrane, the membranes were hybridized with the galectin-3 probe.

affinity with IgE (33). Human galectin-3 possesses an open reading frame of 750 base pairs encoding a 250-amino-acid protein. The N-terminal 41 amino acids are homologous to a serum response transcription factor, and the adjacent sequence (aa 42–106) contains the repetitive proline–glycine-rich motif. The C-terminal domain contains CRD, a sequence motif shared by other S-type lectins. This domain is responsible for the binding, and is well conserved among different species of mammals (34). Although all S-type lectins bind β-galactoside, the specificity of each lectin seems to be unique (13, 35). Our results confirmed the ability of galectin-3 to bind the blood group structures (13, 14).

We examined the selective enrichment of phages expressing galectin-3 fusion proteins during biopanning by two methods; the Southern hybridization and the plaque hybridization. DNA was transferred from the agarose gel, shown in Fig. 1A, onto a Nylon membrane, and the membrane was hybridized with a radiolabeled galectin-3 cDNA probe. The result is shown in Fig. 1B. Five bands (one, three, and one from selections with 50 μg, 5 μg, and 500 ng of target ligand, respectively) were hybridized with the probe. Lysates, containing selected populations of phages, were plated at a density of approximately 500 phages per plate. DNA was transferred onto Nylon membranes, and these membranes were used for the plaque hybridization with the galectin-3 probe. Results from the plaque hybridization experiments are shown in Fig. 3. The numbers of the hybridized plaques were counted, enrichment factors were calculated, and the results are shown in Table 1. These two hybridization experiments clearly demonstrated the selective amplification of phages that expressed galectin-3 structures. The results also showed that the selection was target concentration dependent. Although the enrichment was observed even when the bait ligand was used in lower concentrations, the sequences from the enriched phages were not related to the galectin-3 sequence.

DISCUSSION

We obtained the results that demonstrate the feasibility of phage display technology for the purposes of

TABLE 1

Enrichment of Phages That Express Galectin-3 Fusion Proteins

Target ligand amount	Round of selections	No. of total plaques	No. of galectin-3 positive plaques	Enrichment
Original library	None	8122	1	1
5 ng	1st	412	0	—
	2nd	642	0	—
	3rd	602	0	—
50 ng	1st	502	0	—
	2nd	519	0	—
	3rd	493	0	—
500 ng	1st	301	0	—
	2nd	391	2	41.6
	3rd	625	247	3210.4
5 μg	1st	610	20	266.3
	2nd	650	267	3336.9
	3rd	610	525	6991.5
50 μg	1st	374	16	347.5
	2nd	312	83	2161.1
	3rd	536	359	5440.9

Note. Enrichment of phages expressing galectin-3 fusion protein was examined by plaque hybridization of phages from each round of selection.

detecting carbohydrate-protein interaction and cloning phage particles displaying fusion proteins with carbohydrate affinity. The experiments were originally intended to evaluate the capability of cloning cDNAs encoding glycosyltransferases using this novel technology. We examined whether the phages that expressed A transferase fusion protein could be enriched by serial rounds of biopannings to the target containing blood group H-specific oligosaccharide structures, the substrates for this enzyme. Rather than A transferase, we observed enrichment of phages that expressed galectin-3 fusion proteins after the selections with relatively higher concentrations of target. This enrichment was specific, without doubt, because multiple number of phage clones containing different portions of the galectin-3 were obtained. Enrichment was also observed with lower concentrations of target, however, it remains to be resolved whether or not the enrichment was meaningful because multiple numbers of clones containing different portions of the identical sequences were not obtained in these cases.

Enrichment of phages that express galectin-3 was unexpected. However, considering the reported strong affinity of certain S-type lectins to the blood group structures, this may not be astonishing. Nonetheless, the phage display *cDNA* cloning was shown to be successful in targeting at the proteins with carbohydrate affinity. The same approach may be useful to clone cDNAs that encode other carbohydrate-binding proteins, using the appropriate target ligands for selections. If the binding/washing conditions are optimized, cDNA cloning of proteins with lower affinity may be possible. Various modifications of conditions may be workable. The temperature for phage binding can be changed. Concentration of calcium and other cations may be critical for those proteins that depend on these ions for binding. Phage display system allows both evaluation of the presence of other types of lectins and the determination of their identity. Different from glycosyltransferases, however, some lectins, such as selectins, possess CRD at the N-terminus. Therefore, use of two complementary vector systems that permit the fusion at both the N- and the C-terminus may become necessary.

Selection of glycosyltransferase cDNAs by phage display approach seems to be difficult, even if not impossible, especially when competitive phages that exhibit fusion proteins, such as lectins, with stronger binding properties to the same target are present in the library. The messages for lectins are generally more abundant than those for glycosyltransferases. Therefore, the source of RNA for the library construction may be crucial. Human A transferase was previously purified from lungs to the homogeneity through several steps of procedures including Triton X-100 extraction, Sepharose 4B (UDP) chromatography, cation-exchange chromatography, and reverse phase chromatography (36). Because we were unsure whether the phages would tolerate these procedures and because we had crude mucin fraction with blood group H activity, we took a shortcut approach by utilizing the affinity of the enzyme to the substrate. Chromatographic steps similar to some of those used in enzyme purification may be applied to isolate the phages displaying A transferase. Use of high copy number display T7Select 415b Vector may increase the valence and enhance the phage's affinity with the ligand. This vector, however, can only incorporate a maximum of 39 amino acids, which may be insufficient to code for CRD when taking into account the general belief that about 100 amino acid residues are necessary to form a self-contained sugar-binding site (12). Future success of cloning glycosyltransferase cDNAs by phage display, therefore, may heavily depend on the development of new vectors that display multiple copies of long protein sequences.

Lectins are involved in pathological as well as physiological phenomena, and may also play an important role in progression and metastasis of cancer. We have shown that this novel approach of phage display is useful in identifying proteins with carbohydrate-binding properties. Using this technique, new types of lectins, if any, may be identified, CRD epitopes may be mapped, and our understanding of the structures and functions of these molecules involved in carbohydrate-protein interactions may be greatly advanced.

ACKNOWLEDGMENTS

The authors thank Dr. Manuel Perucho for general support and encouragement. We are grateful to Ami Yamamoto for editing the manuscript. This work has been supported in part by DoD BCRP Grant DAMD17-98-1-8168 (to F.Y.) and by Kyowa Hakko Kogyo Co. (to F.Y.).

REFERENCES

1. Hakomori, S. (1981) *Annu. Rev. Biochem.* **50,** 733–764.
2. Feizi, T. (1985) *Nature* **314,** 53–57.
3. Springer, T. A. (1994) *Cell* **76,** 301–314.
4. Lowe, J. B., Stoolman, L. M., Nair, R. P., Larsen, R. D., Berhend, T. L., and Marks, R. M. (1990) *Cell* **63,** 475–484.
5. Phillips, M. L., Nudelman, E., Gaeta, F. C. A., Perez, M., Singhal, A. K., Hakomori, S-I., and Paulson, J. C. (1990) *Science* **250,** 1130–1132.
6. Walz, G., Aruffo, A., Kolanus, W., Bevilacqua, M., and Seed, B. (1990) *Science* **250,** 1132–1135.
7. Lawrence, M. B., and Springer, T. A. (1991) *Cell* **65,** 859–873.
8. Ley, K., Gaehtgens, P., Fennie, C., Singer, M. S., Lasky, L. A., and Rosen, S. D. (1991) *Blood* **77,** 2553–2555.
9. Bevilacqua, M. P., Stengelin, S., Gimbrone, M. A., Jr., and Seed, B. (1989) *Science* **243,** 1160–1165.
10. Nakamori, S., Kameyama, M., Imaoka, S., Furukawa, H., Ishikawa, O., Sasaki, Y., Kabuto, T., Iwanaga, T., Matsushita, Y., and Irimura, T. (1993) *Cancer Res.* **53,** 3632–3637.
11. Barondes, S. H. (1984) *Science* **223,** 1259–1264.
12. Drickamer, K. (1994) *in* Molecular Glycobiology (Fukuda, M., and Hindsgaul, O., Eds.), pp. 53–87, IRL Press, Oxford, UK.
13. Abbott, M. W., Hounsell, E. F., and Feizi, T. (1988) *Biochem. J.* **252,** 283–287.
14. Sato, S., and Hughes, R. C. (1992) *J. Biol. Chem.* **267,** 6983–6990.
15. Cherayil, B. J., Chaitovitz, S., Wong, C., and Pilai, S. (1990) *Proc. Natl. Acad. Sci USA* **87,** 7324–7328.
16. Lotz, M. M., Andrews, C. W., Korzelius, C. A., Lee, E. C., Steele, G. D., Clarke, A., and Mercurio, A. M. (1993) *Proc. Natl. Acad. Sci USA* **90,** 3466–3470.
17. Raz, A., Meromsky, L., and Lotan, R. (1986) *Cancer Res.* **46,** 3667–3672.
18. Perillo, N. L., Pace, K. E., Seihamer, J. J., and Baum, L. G. (1995) *Nature* **378,** 736–739.
19. Wells, V., and Mallucci, L. (1991) *Cell* **64,** 91–97.
20. Yamamoto, F., Marken, J., Tsuji, T., White, T., Clausen, H., and Hakomori, S. (1990) *J. Biol. Chem.* **265,** 1146–1151.
21. Yamamoto, F., Clausen, H., White, T., Marken, J., and Hakomori, S. (1990) *Nature* **345,** 229–233.
22. Yamamoto, F., and Hakomori, S. (1990) *J. Biol. Chem.* **265,** 19257–19262.
23. Smith, G. P (1985) *Science* **228,** 1315–1317.
24. Winter, G., and Milstein, C. (1991) *Nature* **349,** 293–299.
25. Clackson, T., and Hoogenboom, H. R., Griffiths, A. D., and Winter, G. (1991) *Nature* **352,** 624–628.
26. Wrighton, N. C., Farrell, F. X., Chang, R., Kashyap, A. K., Barbone, F. P., Mulcahy, L. S., Johnson, D. L., Barrett, R. W., Jolliffe, L. K., and Dower, W. J. (1996) *Science.* **273,** 458–464.
27. Phage Display of Peptides and Proteins: A Laboratory Manual (1996) (Kay, B., *et al.,* Eds.), Academic Press, San Diego.
28. Pasqualini, R., Koivunen, E., and Ruoslahti, E. (1995) *J. Cell Biol.* **130,** 1189–1196.
29. Pasqualini, R., and Ruoslahti, E. (1996) *Nature* **380,** 364–366.
30. Kuwabara, I., Maruyama, H., Mikawa, Y. G., Zuberi, R. I., Liu, F-T., and Maruyama, I. N. (1997) *Nature Biotechnol.* **15,** 74–78.
31. Kabat, A. (1956) Blood Group Substances, Academic Press, New York.
32. Yamamoto, F., McNeill, P. D., and Hakomori, S. (1992) *Biochem. Biophys. Res. Commun.* **187,** 366–374.
33. Albrandt, K., Orida, N. K., and Liu, F. T. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 6859–6863.
34. Robertson, M. W., Albrandt, K. A., Keller, D., and Liu, F-T. (1990) *Biochemistry* **29,** 8093–8100.
35. Leffler, H., and Barondes, S. H. (1986) *J. Biol. Chem.* **261,** 10119–10126.
36. Clausen, H., White, T., Takio, K., Titani, K., Stroud, M., Holmes, E., Karkov, J., Thim, L., and Hakomori, S. (1990) *J. Biol. Chem.* **265,** 1139–1145.